# Temporal Aggregation and Causality in Multiple Time Series Models*

Jörg Breitung
Humboldt University

Norman R. Swanson
Pennsylvania State University

March 10, 1998

**Abstract**

In this paper we characterize what has sometimes been referred to in the literature as instantaneous causality, by examining the consequences of temporal aggregation in (possibly) Granger causal systems of variables. Our approach is to compare the concept of contemporaneous correlation due to Swanson and Granger (1997) with that of Granger causality. Using asymptotic theory based on large aggregation intervals we derive conditions for a correspondence between both concepts. These results allow us to differentiate between spurious contemporaneous correlation arising because of aggregation, and true Granger causality. Monte Carlo experiments indicate that the asymptotic results provide a reliable guidance for finite samples and finite aggregation intervals.

JEL Classification: C32, C43, C51.

KEYWORDS: instantaneous causality, Granger causality, contemporaneous correlation, temporal aggregation, stock and flow variables.

---

* *Corresponding Author*: Joerg Breitung, Institute of Statistics and Econometrics, Humboldt University, Berlin, Spandauer Str. 1, D-10178 Berlin, Germany. The authors wish to thank Clive W.J. Granger, Helmut Lütkepohl and Michael Neumann for helpful comments. The first author gratefully acknowledges financial support from the *Sonderforschungsbereich 373* at the Humboldt University Berlin. The second author thanks the Research and Graduate Studies Office at Pennsylvania State University for research support.

# 1 Introduction

Aggregation poses many interesting questions which have been explored in time series analysis and which yet remain to be explored. For example, models which encorporate linear difference equations are often justified on the basis of some Granger causal ordering. However, when series are not observed at what might be termed their *natural* frequency, corresponding to the causal structure which the modeler has in mind, difficulties in determining whether the posited causal structure of the model is valid may arise. In particular, if a Granger causal structure is associated with a *natural* frequency of observation, which is much lower than the observed frequency of observation, then the *true* Granger causal relationship among the data might rather appear in the form of contemporaneous correlation among the variables in the model. Indeed a rich literature which tackles such issues has accumulated over many years. An early example is Quenouille (1958), where it is shown that autoregressive moving average processes of order $(p, q \leq p)$ remain the same order when observations are sampled at $k$ times their *natural* frequency. Amemiya and Wu (1972), and Brewer (1973) refine and generalize Quenouille's result by including "exogenous" regressors. Zellner and Montmarquette (1971) discuss the effects of temporal aggregation on estimation and testing. Engle (1969) and Wei (1978) analyze the effects of temporal aggregation on parameter estimation in a distributed lag model. Granger (1987) discusses the implications of aggregation on systems with common factors. Other important contributions in the area of temporal aggregation are Stram and Wei (1986), Lütkepohl (1987), and Marcellino (1996), to name but a few.

Weiss (1984) uses the method of Wei (1981) to discuss the effects of systematic sampling and temporal aggregation on ARMA and ARMAX models. He notes that *'Some care needs to be taken in causality testing as causality is defined for the true data-generating process and not for the aggregated data.'* In this paper we show that Weiss' (1984) comment is quite true. We do this by characterizing what has sometimes been referred to in the literature as *instantaneous* causality, and examining the consequences of temporal aggregation in (possibly) Granger causal systems of variables.

In particular, our approach is to compare the concept of contemporaneous correlation due to Swanson and Granger (1997) with that of Granger causality. Using asymptotic theory based on large aggregation intervals we derive conditions for a correspondence between both concepts. These results allow us to differentiate between spurious contemporaneous correlation arising because of aggregation, and true Granger causality.

The rest of the paper is organized as follows. In Section 2, we review the concepts of Granger causality and contemporaneous correlation, pointing out, for example, that contemporaneous correlation is sometimes equated with instantaneous causation. We do not, though, argue for the use of *instantaneous causation* as a concept in economics, but rather show that contemporaneous correlation can arise when aggregated data are examined, even if the same data exhibit no contemporaneous correlation at higher frequencies of observation. In Section 3, we provide a framework for the examination of the effects of temporal aggregation on (Granger) causal inference, while Section 4 gives related asymptotic results. Section 5 outlines our findings concerning the relationship between contemporaneous correlation (which we in some cases use to define what we term *contemporaneous causality*) and Granger causality. Section 6 provides Monte Carlo evidence based on our asymptotic findings, and Section 7 provides concluding remarks.

# 2 Granger and Contemporaneous (Instantaneous) Causality

Following Granger (1969), consider a conditional distribution with respect to two information sets which are available at time $t$, say $\mathcal{I}'_t$ and $\mathcal{I}_t = \{\mathcal{I}'_t, y_{j,t}, y_{j,t-1}, \ldots\}$. In the following, we use a conditional mean definition of causality. Specifically, we define a variable $y_{j,t}$ to be Granger causal for the variable $y_{i,t}$ if

$$E(y_{i,t+1}|\mathcal{I}_t) \neq E(y_{i,t+1}|\mathcal{I}'_t). \tag{1}$$

If $y_{j,t}$ is causal for $y_{i,t}$ we write $y_{j,t} \to y_{i,t}$. When one variable is *not* Granger causal for another, we write $y_{i,t} \not\to y_{j,t}$. It is often found in empirical examinations of economic data, though, that Granger causality runs in both directions. In such cases, we write $y_{i,t} \leftrightarrow y_{j,t}$.

Given these definitions, a natural question to ask is whether or not contemporaneous correlation between different variables should be equated with some sort of causality definition. This topic has been addressed in a number of papers including Pierce and Haugh (1977), Granger (1988), Swanson and Granger (1997), and the references therein. Pierce and Haugh (1977) propose a definition of *instantaneous* causality, where one variable, say $y_{j,t}$, *instantaneously* causes another variable (in mean), say $y_{i,t}$, if the contemporaneous values of $y_{j,t}$ are useful for "forecasting" contemporaneous values of $y_{i,t}$. A necessary condition for linear processes is that $\rho(y_{i,t}, y_{j,t}) \neq 0$, where $\rho(\cdot)$ denotes the simple correlation between $y_{i,t}$ and $y_{j,t}$. Obviously, there is a symmetry problem with the definition, as no direction of the relationship between $y_{i,t}$ and $y_{j,t}$ can be deduced. Further, as pointed out by Granger (1988), *apparent instantaneous* causality may result from (i) a time aggregation of the data, given that the original time interval over which the data are generated is much less than the interval over which the data are measured, and (ii) omitted or unobserved variables may lead to causality among a group of variables, thus leading to contemporaneous correlation in some situations (Lütkepohl (1982)). As mentioned above, in this paper we examine the effects of temporal aggregation within the context of vector autoregressive (VAR) models.

In addition to the obvious need to understand the consequences of temporal aggregation in time series, another reason for examing the effects of temporal aggregation on VAR models is pointed out by Swanson and Granger (*henceforth* SG (1997)), who examine partial covariances (or partial correlations) defined as $\gamma(r, s|u) = E(r_u, s_u)$ with $r_u = r - E(r|u)$ and $s_u = s - E(s|u)$ within the context of choosing causal orderings, Wold causal chains, or Choleski decompositions with which to orthogonalize the errors (or residuals) in a VAR model. Consider conditions of the form

$$E(y_{j,t+1}|y_{i,t+1}, z_{t+1}, \mathcal{I}_t) = E(y_{j,t+1}|z_{t+1}, \mathcal{I}_t), \qquad (2)$$

where the vector $z_{t+1} = \{y_{l,t+1}|l \neq i, j\}$ comprises all variables except $y_{i,t+1}$ and

$y_{j,t+1}$. Now, note that necessary and sufficient conditions for (2) to hold can be given in terms of partial covariance (correlation) restrictions of the form:

$$\gamma(y_{i,t+1}, y_{j,t+1} | z_{t+1}, \mathcal{I}_t) = 0. \tag{3}$$

These sorts of restrictions arise in an obvious manner in restricted and unrestricted recursive models. For example, consider the model

$$\begin{aligned} y_{1,t} &= u_{1,t} && (4) \\ y_{2,t} &= a y_{1,t} + u_{2,t} && (5) \\ y_{3,t} &= b y_{2,t} + u_{3,t}, && (6) \end{aligned}$$

where $b \neq 0$ and the $u_{j,t}$ are mutually uncorrelated white noise random variables with $E(u_{j,t}) = 0$, $(j = 1, 2, 3)$. Using the fact that $\gamma(r, s|s) = 0$, and assuming that $E(r) = 0$, it follows from (6) that:

$$\gamma(y_{1,t}, y_{3,t}|y_{2,t}) = b\ \gamma(y_{1,t}, y_{2,t}|y_{2,t}) + \gamma(y_{1,t}, u_{3,t}|y_{2,t}) = 0.$$

Accordingly, in the model given by equations (4) – (6) there is a contemporaneous relationship between $y_{2,t}$ and $y_{3,t}$, and between $y_{1,t}$ and $y_{2,t}$, but not between $y_{1,t}$ and $y_{3,t}$, when conditioning on the remaining variable in the system, so that restrictions of the form given by (3) are indeed implied by the model. One way to express this result is by using a directed acyclic graph (DAG).[1] In this example, the DAG is equivalent to the Wold causal chain described in Lütkepohl (1991), and used to *orthogonalize* the errors in a VAR model. Further, it should be noted that the DAG is solely meant to be a convenient way of representing the system, (4) – (6), and is written as $y_{1,t} \Rightarrow y_{2,t} \Rightarrow y_{3,t}$ in this particular example. Whenever one variable (say $y_{2,t}$) operates a *cut* between $y_{1,t}$ and $y_{3,t}$, for example, then there is no contemporaneous correlation between $y_{1,t}$ and $y_{3,t}$, conditional on $y_{2,t}$. On the other hand, $y_{1,t}$ and $y_{2,t}$ ($y_{2,t}$ and $y_{3,t}$) *are* contemporaneously correlated when conditioned on $y_{3,t}$ ($y_{1,t}$).

In general, graphs of this sort provide a useful shortcut for expressing the contemporaneous relationships among variables. Swanson and Granger (1997)

[1] See, e.g., Glymour, Scheines, Spirtes, and Kelly (1987) and Swanson and Granger (1997) for further details.

formalize the use of partial correlations and DAGs for orthogonalizing the errors in VAR models, by providing a simple regression based technique for testing the adequacy of some given recursive structural model of the errors, thus avoiding the well known "pitfall" of having to arbitrarily choose the ordering of the variables (and hence errors) before constructing impulse response functions.

However, SG do not discuss in detail how contemporaneous correlations of the type which they examine might arise in practice and indeed that simple recursive VAR models which exhibit uni-directional Granger causality *and* no contemporaneous correlation (sometimes called *instantaneous causality*), are characterized by increasing levels of contemporaneous correlation when the data are temporally aggregated.

# 3 Causal Inference in Aggregated Time Series

In order to examine the relationship between contemporaneous correlation and Granger causality in the presence of temporally aggregated data, we consider two procedures which are used for the temporal aggregation of economic time series (e.g. see Lütkepohl (1987)). For flow data, time series values are cumulated (or averaged) at $k$ successive time periods

$$\bar{y}_t = \sum_{j=0}^{k-1} y_{t-j}$$

and the aggregated series results from applying *skip*-sampling of the form

$$\bar{Y}_T = \bar{y}_{kT} \quad T = 1, 2, \ldots,$$

where it is assumed that the time series starts at the beginning of the aggregation period. Stock data are aggregated by directly applying the skip-sampling scheme to the data, so that $Y_T = y_{kT}$ for $T = 1, 2, \ldots$.

Now, assume that the $d^{th}$ differences of an $m \times 1$ vector of time series $\{y_t\}_1^T$ is stationary with a Wold representation given by:

$$\begin{aligned} \Delta^d y_t &= \varepsilon_t + C_1\varepsilon_{t-1} + C_2\varepsilon_{t-2} + C_3\varepsilon_{t-3} + \cdots \\ &= C(L)\varepsilon_t, \end{aligned} \tag{7}$$

where $L$ is the lag operator defined as $y_{t-k} = L^k y_t$, $\Delta = 1 - L$, and $\varepsilon_t$ is assumed to be an i.i.d. vector process with $E(\varepsilon_t) = 0$ and $E(\varepsilon_t \varepsilon_t') = \Omega$. Without loss of generality we ignore deterministic terms (e.g. constants and linear deterministic trends). Further, we confine our attention to cases where $d = 0$ and $d = 1$. If the moving average (MA) polynomial, $C(L)$, is invertible[2], there exists a VAR representation of the form:

$$\Delta^d y_t = A_1 \Delta^d y_{t-1} + A_2 \Delta^d y_{t-2} + \cdots + A_p \Delta^d y_{t-p} + \varepsilon_t \ , \tag{8}$$

where the lag order $p$ may be infinite.

If $y_t$ is a stationary VARMA process, then the aggregated vectors $Y_T$ or $\bar{Y}_T$ are VARMA processes as well (see e.g. Lütkepohl (1987)). Provided that the MA polynomial is invertible, the aggregated vector of stock variables has an (infinite) VAR representation given by:

$$Y_T = A_1^{(k)} Y_{T-1} + A_2^{(k)} Y_{T-2} + \cdots + U_T$$

where $E(U_T U_T') := \Omega^{(k)}$. Note that our above notation suppresses the dependence of $Y_T$ and $U_T$ on $k$. A similar representation can be given for aggregated flow variables. In what follows we confine our analysis to the case of a stationary vector of stock variables. Flow variables and nonstationary variables can be treated in analogous fashion.

Granger causality of the form $y_{i,t} \to y_{j,t}$ (see equation (1)) arises if

$$e_i' A_\tau e_j \neq 0 \quad \text{for some } \tau \in \{1, 2, \ldots\},$$

where $e_i$ ($e_j$) is the $i^{th}$ ($j^{th}$) column vector of identity matrix $I_m$. Further, the definitions of contemporaneous correlation given by (2) and (3) imply restrictions on the transformed VAR representation of the aggregated data:

$$Y_T = B_0^{(k)} Y_T + B_1^{(k)} Y_{T-1} + B_2^{(k)} Y_{T-2} + \cdots + \eta_T \ , \tag{9}$$

where the diagonal elements of $B_0$ are zero and $\eta_T = (I - B_0) U_T$ is a vector of mutually uncorrelated errors. This representation results from the multiplication

[2]The case of a system of cointegrated variables is considered below.

of (8) by $(I - B_0)$, where $B_0$ is chosen such that $(I - B_0)\Omega^{(k)}(I - B_0)'$ is a diagonal matrix, and the components of $\eta_t$ are mutually uncorrelated, so that standard impulse response functions can be constructed, for example. According to (2) and (3), $Y_{i,T}$ is contemporaneously correlated with $Y_{j,T}$ if

$$e_i' B_0^{(k)} e_j \neq 0 \quad \text{for } i \neq j.$$

At this juncture, it is useful to provide a definition which links the concepts of Ganger causality and contemporaneous correlation which we have discussed thus far.

**Definition 1:** *Let $Z_{T+1} = \{Y_{l,T+1} | l \neq i, j\}$ be a $(m-2) \times 1$ vector of aggregated flow or stock variables, except $Y_{i,T+1}$ and $Y_{j,T+1}$, where the elements of $Y_{T+1}$ are stacked. Further, define the information sets: $\mathcal{J}_T = \{Y_T, Y_{T-1}, Y_{T-2}, \ldots\}$ and $\mathcal{J}_T' = \{Z_T, Y_{j,T}, Z_{T-1}, Y_{j,T-1}, \ldots\} = \mathcal{J}_T - \{Y_{i,N} | N \leq T\}$. Then $Y_{i,T}$ is said to be a cause of $Y_{j,T}$ with respect to the associated aggregation interval $k$, if*

$$E(Y_{j,T+1} | Y_{i,T+1}, Z_{T+1}, \mathcal{J}_T) \neq E(Y_{j,T+1} | Z_{T+1}, \mathcal{J}_T').$$

For equation (9), this definition implies that $Y_{i,T}$ is a cause for $Y_{j,T}$ with respect to the aggregation level $k$, if:

$$e_i' B_\tau^{(k)} e_j \neq 0 \quad \text{for some } \tau \in \{0, 1, 2, \ldots\}$$

It follows that *for finite $k$*, contemporaneous correlation according to equation (2) is sufficient (but not necessary) for causality according to Definition 1. For $k = 1$, the process is observed at its original time scale. This time scale is defined as the time resolution where cause occurs before effect. As argued by Granger (1988) and Swanson and Granger (1997), a reasonable assumption is that at such time scale there is no contemporaneous correlation between the variables. This implies that $\Omega$ is a diagonal matrix. Accordingly, $B_0$ is a zero matrix and causality according to Definition 1 coincides with Granger causality. On the other hand, if $k$ tends to infinity, our definition of causality is equivalent to what we have thus far termed contemporaneous correlation, and what we will also refer to as *contemporaneous causality*. Hence, our definition of causality has the property that it corresponds to Granger causality for $k = 1$ and corresponds to contemporaneous (or perhaps, *instantaneous*) causality for $k \to \infty$.

Using the above framework, it is possible to investigate the relationship between the traditional concept of Granger causality and the concept of contemporaneous causality. Since it seems unlikely that a process will be observed with a sampling frequency such that the cause strictly occurs before the effect (i.e. $k = 1$), in practice, it is interesting to determine whether or not it is possible to identify *original* causal patterns from aggregated data, given the presence of certain contemporaneous causation patterns.

# 4 Asymptotic Results

In this section, we examine the case for which the time lag between *cause* and *effect* is small relative to the sample frequency. In other words, we assume that the aggregation interval, $k$, tends to infinity but as a slower rate than $T$, i.e., $k/T \to 0$. A related framework is used by Christiano and Eichenbaum (1987). They assume the existence of a continuous time version of Wold's decomposition theorem such that $y(t) = \int f(\tau)\varepsilon(t-\tau)\mathrm{d}\tau$. The actual time series vector is observed at integer values of $t$ so that $y_t = y(t)$ for $t = 1, 2, \ldots$. The asympotic results presented in this section remain substantially the same when assuming a continous time process instead of a discrete process.

We begin by assuming that the covariance matrix $\Omega$ is diagonal. Given that the notion of Granger causality implicitly assumes that a *cause* is strictly prior to an *effect* (in a time series sense), one may imagine a sufficiently fine time resolution (e.g. a high enough frequency of data), whereby cause is indeed strictly prior to effect, in which case $\Omega$ will be diagonal. The following proposition summarizes some properties of the limiting process for an aggregated vector time series, say $y_t$, as $k \to \infty$.

**Proposition 1:** *Let $y_t$ be generated by an $m$ dimensional linear process $y_t =$*

*$\varepsilon_t + C_1\varepsilon_{t-1} + C_2\varepsilon_{t-2} + \cdots$, where $E(\varepsilon_t\varepsilon_t') = \Omega$, and $y_t$ is one-summable such that $\sum_{j=1}^{\infty} j|C_j| < \infty$, where $|C| = max_{i,j}|C_{i,j}|$. As $k \to \infty$, the processes for the aggregated vectors $Y_T$ and $\bar{Y}_T$ have the properties that:*

**stock variables:**

$$(i) \qquad \lim_{k\to\infty} E(Y_T Y_T') = \left(I_m + \sum_{i=1}^{\infty} C_i\right) \Omega \left(I_m + \sum_{i=1}^{\infty} C_i'\right)$$

$$(ii) \qquad \lim_{k\to\infty} E(Y_T Y_{T+j}') = 0 \quad \text{for } j \geq 1,$$

**flow variables:**

$$(iii) \qquad \lim_{k\to\infty} \frac{1}{k} E(\bar{Y}_T \bar{Y}_T') = 2\pi f_y(0)$$

$$(iv) \qquad \lim_{k\to\infty} E(\bar{Y}_T \bar{Y}_{T+1}') = \sum_{j=1}^{\infty} \left(I_m + \sum_{i=1}^{j} C_i\right) \Omega \left(\sum_{i=j+1}^{\infty} C_i\right)'$$

$$(v) \qquad \lim_{k\to\infty} E(\bar{Y}_T \bar{Y}_{T+j}') = 0 \quad \text{for } j \geq 2,$$

*where $f_y(\omega)$ denotes the spectral density matrix of $y_t$ at frequency $\omega$.*

According to Proposition 1, it turns out that for $k \to \infty$, the aggregated processes are asymptotically white noise. Of course this result is not particularly surprising, since it is intuitively plausible that with increasing sampling interval, short-run dynamics disappear. Note also that for the aggregation of a flow variable the process must be divided by the square root of $k$ in order to obtain a finite variance. Furthermore, for moderate $k$ it is expected that aggregated flow variables are well approximated by a vector MA(1) process. The reason for this is that according to (iv), the first order autocorrelation is $O(k)$, while (v) implies that higher order autocorrelations are $o(1)$.

Next, assume that $y_t$ is a vector of integrated variables such that $\Delta y_t$ is stationary with a Wold representation as in (7). Further, assume that the matrix $\bar{C} = I_m + \sum_{j=1}^{\infty} C_j$ is of full rank (i.e. the variables in $y_t$ are not cointegrated).

**Proposition 2:** *Let $\Delta y_t$ be generated by an $m$ dimensional linear process $y_t = \varepsilon_t + C_1\varepsilon_{t-1} + C_2\varepsilon_{t-2} + \cdots$, where it is assumed that $E(\varepsilon_t\varepsilon_t') = \Omega$ and $\sum_{j=1}^{\infty} j|C_j| < \infty$. As $k \to \infty$, the processes for the aggregated vectors $Y_T$ and*

$\bar{Y}_T$ *are characterized by:*

$$
\begin{array}{ll}
 & \textbf{stock variables:} \\
(i) & \lim_{k\to\infty} \frac{1}{k} E(Y_T - Y_{T-1})(Y_T - Y_{T-1})' = 2\pi f_{\Delta y}(0) \\
(ii) & \lim_{k\to\infty} \frac{1}{k} E(Y_T - Y_{T-1})(Y_{T+j} - Y_{T+j-1})' = 0 \quad \text{for } j \geq 1, \\
 & \\
 & \textbf{flow variables:} \\
(iii) & \lim_{k\to\infty} \frac{1}{k^3} E(\bar{Y}_T - \bar{Y}_{T-1})(\bar{Y}_T' - \bar{Y}_{T-1})' = \frac{4\pi}{3} f_{\Delta y}(0) \\
(iv) & \lim_{k\to\infty} \frac{1}{k^3} E(\bar{Y}_T - \bar{Y}_{T-1})(\bar{Y}_{T+1} - \bar{Y}_T)' = \frac{\pi}{3} f_{\Delta y}(0) \\
(v) & \lim_{k\to\infty} \frac{1}{k^3} E(\bar{Y}_T - \bar{Y}_{T-1})(\bar{Y}_{T+j} - Y_{T+j-1})' = 0 \quad \text{for } j \geq 2,
\end{array}
$$

*where* $f_{\Delta y}(\omega)$ *denotes the spectral density matrix of* $\Delta y_t$ *at frequency* $\omega$.

Based on Proposition 2, it thus follows that as $k$ tends to infinity, the vector of aggregated flow variables has the vector MA(1) representation:

$$\frac{1}{\sqrt{k^3}}(\bar{Y}_T - \bar{Y}_{T-1}) = U_T + (2 - \sqrt{3})U_{T-1}, \tag{10}$$

where

$$E(U_T U_T') = \frac{2\pi}{1 + (2 - \sqrt{3})^2} f_{\Delta y}(0) \ .$$

Note that for the special case where $m = 1$ (a single time series), our results correspond to the result of Working (1960) who shows that the first order autocorrelation of the increments from an aggregated random walk is 0.25.

The asymptotic results of Proposition 1 and 2 imply that for difference stationary stock variables as well as stationary and difference stationary flow variables the contemporaneous relationship of the limiting process reflect the causal relationship at frequency zero in the sense of Geweke (1986) and Granger and Lin (1995).

It is interesting to consider the case where $y_t$ is cointegrated. In this case we assume that there exists a vector error correction representation of the form:

$$\Delta y_t = \Pi y_{t-1} + \Delta y_{t-1} + A_2 \Delta y_{t-2} + \cdots + A_p \Delta y_{t-p} + \varepsilon_t, \tag{11}$$

where the matrix $\Pi = \alpha\beta'$ is of rank $0 < r < m$ and where $\alpha$ and $\beta$ are $m \times r$ matrices. We rule out the case where $y_t$ is integrated of order two or higher, and assume that $\Delta y_t$ and $\beta' y_t$ are I(0), using the terminology of Engle and Granger (1987).

In order to discuss the effects of aggregation in the context of cointegrated variables, it is useful to define the matrix $Q' = [\beta, \gamma]^{-1}$, where $\beta$ is a $n \times r$ matrix of cointegration vectors such that $z_t = \beta' y_t$ is I(0). The matrix $\gamma$ is some $n \times (n-r)$ matrix, which is linearly independent of $\beta$. The linear combinations $w_t = \gamma' y_t$ are assumed to be $I(1)$. From Proposition 1 (iii) it follows that the variance of the aggregated vector of variables, $\bar{Z}_T - \bar{Z}_{T-1}$, is $O(k)$ while the variance of $\bar{W}_T - \bar{W}_{T-1}$ is $O(k^3)$. Hence, as the aggregation interval $k$ tends to infinity, the variance of the "nonstationary linear combinations" dominates the variance of the "error correction terms". Consequently,

$$\begin{aligned} \frac{1}{k^{3/2}}(\bar{Y}_T - \bar{Y}_{T-1}) &= \frac{1}{k^{3/2}} Q \begin{bmatrix} \bar{Z}_T - \bar{Z}_{T-1} \\ \bar{W}_T - \bar{W}_{T-1} \end{bmatrix} \\ &= \frac{1}{k^{3/2}} Q_2 (\bar{W}_T - \bar{W}_{T-1}) + O_p(k^{-1}), \end{aligned}$$

where $Q_2$ is the lower $m \times (m-r)$ block of $Q$. This implies that the differences of $\bar{Y}_T$ possess a singular distribution as $k$ tends to infinity. (It is important to note that the limiting processes of the aggregated variables have a singular spectral density matrix for all frequencies $0 \leq \omega \leq \pi$, while the spectral density matrix of $y_t$ is singular at $\omega = 0$ only.) In other words, the limiting behavior of the aggregated time series is dominated by the stochastic trends in the system, and thus, the standardized variance of the error correction terms tends to zero. Since this does not seem to be a relevant feature of observed time series, for present, we exclude the aggregation of cointegrated variables from our analysis, leaving further examination of aggregated cointegrated variables to future research.

# 5 The Relationship between Granger and Contemporaneous Causality

Using the causality definition given in Section 2, we are able to consider the relationship between Granger causality ($k = 1$) and contemporaneous causality ($k \to \infty$). As mentioned above, this comparison is of interest, as it is unlikely that any given group of variables will be observed with a time scale which allows an unambiguous ordering of *cause* and *effect*, in time. However, we cannot reasonably hope that any correspondence between our two varieties of causality will be one-to-one, given that aggregation imparts a loss of information on our system. Specifically, as different causal structures may have the same limiting distributions, we cannot uniquely identify the underlying causal structure by considering the limiting process alone. An obvious example of this problem is the difficulty of identifying the direction of the Granger causal relationship from contemporaneous correlations alone. Nevertheless, it is possible to derive some general results which characterize the relationship between the causal structure of original and aggregated processes. In particuar, the following proposition gives sufficient conditions for non-causality between two aggregated variables.

**Proposition 3:** *Assume that either:*

| | |
|---|---|
| *(i)* | *$y_t$ is a vector of stationary flow variables, or* |
| *(ii)* | *$y_t$ is a vector of difference stationary flow variables, or* |
| *(iii)* | *$y_t$ is a vector of difference stationary stock variables.* |

*If there is no Granger causality between $y_{i,t}$ and $y_{j,t}$ and*

$$(a) \quad y_{i,t} \not\to y_{l,t} \quad \text{or} \quad (b) \quad y_{j,t} \not\to y_{l,t} \quad \text{for all } l \neq i, j,$$

*then, as $k \to \infty$, we have for the partial correlations of the aggregated variables that:*

$$\text{for case } (i) \qquad \rho(\bar{Y}_{i,T}, \bar{Y}_{j,T} | \bar{Y}_{l,T}) = 0$$

$$\text{for case } (ii) \qquad \rho(\Delta\bar{Y}_{i,T}, \Delta\bar{Y}_{j,T} | \Delta\bar{Y}_{l,T}) = 0$$

$$\text{for case } (iii) \qquad \rho(\Delta Y_{i,T}, \Delta Y_{j,T} | \Delta Y_{l,T}) = 0$$

*for $l \neq i, j$.*

Proposition 3 gives sufficient conditions for ruling out spurious contemporaneous causality. If conditions (a) or (b) are violated it may be the case that there is contemporaneous causality between aggregated variables, although there is no Granger causality at the original time scale. In this case we will speak of *spurious* contemporaneous causality.

Necessary *and* sufficient conditions for Proposition 3 to hold can be obtained from the Choleski decomposition:

$$(I_m - \bar{A})\Omega(I_m - \bar{A}') = RR', \tag{12}$$

where $\bar{A} = \sum_{j=1}^{p} A_j$, $\Omega$ is a diagonal matrix and $R$ is an upper triangular matrix. A necessary and sufficent condition to exclude spurious contemporaneous causality between $\bar{Y}_{1,t}$ and $\bar{Y}_{2,t}$ is that the (1,2) element of $R$ is zero. However, the elements of $R$ are nonlinear functions of the elements of $\bar{A}$, so that this condition cannot be verified without the knowledge of precise parameter values of the process which describes $y_t$. Therefore, the practical value of such a condition is quite limited.

If $y_{i,t} \rightarrow y_{j,t}$ and $y_{j,t} \rightarrow y_{i,t}$ we say that there is feedback causality between $y_{i,t}$ and $y_{j,t}$. An important consequence of Proposition 3 can be derived for the case that there is no feedback causality among the variables.

**Corollary 1:** *For the cases (i) – (iii) of Proposition 3 and under the assumption that there is no feedback causality among the variables it follows that, as $k \rightarrow \infty$, there is no spurious causality among the aggregated variables of the system.*

Whenever there is no feedback Granger causality, the variables of the system can be arranged such that one of the conditions (a) and (b) of Proposition Proposition 3 is satisfied. This rules out the case of spurious contemporary causality.

Another (less trivial) consequence of Proposition 3 can be derived for a trivariate system. Following Dufour and Renault (1998, Definition 2.2) and Lütkepohl and Burda (1997) we say that $y_{j,t}$ does not cause $y_{i,t}$ *at horizon k"* if

$$E(y_{i,t+h}|\mathcal{I}_t) = E(y_{i,t+h}|\mathcal{I}'_t), \tag{13}$$

where $\mathcal{I}_t$ and $\mathcal{I}'_t$ are the same information sets as in (1). Obviously, the usual definition of causality given in (1) is a special case with $h = 1$. If $y_{i,t}$ does not cause $y_{j,t}$ *at any horizon* we write $y_{i,t} \not\to_\infty y_{j,t}$. For a trivarite system the following result holds.

**Corollary 2:** *Let* $\Delta^d y_t = [y_{1,t}, y_{2,t}, y_{3,t}]'$ *be a stationary* $3 \times 1$ *vector with* $d \in \{0, 1\}$ *and invertible MA representation. If* $y_{2,t} \not\to_\infty y_{3,t}$ *and (i) – (iii) of Proposition 3 hold, then as* $k \to \infty$*, there is no spurious causality among the aggregated counterparts.*

This result is intuitively plausible because the assumption of no causality at any horizon rules out indirect causal effects via the remaining variable $y_{1,t}$. Accordingly, for $y_{1,t} \to y_{3,t}$ we must rule out that $y_{2,t}$ causes $y_{1,t}$ at longer lag horizons $h$, since otherwise $y_{2,t}$ may be used to predict $y_{3,t+h}$ via $y_{1,t+j}$, where $0 < j < h$. Unfortunately, we were not able to generalize this result to higher dimensional systems with $M > 3$.

In order to illustrate the results in this section, it is useful to construct some examples based on a trivariate VAR(1) model for $y_t$, where all innovation terms represent mutually uncorrelated white noise processes.

**Example A:** Assume that $y_t$ has a causal structure given by $y_{1,t} \to y_{2,t}$ and $y_{2,t} \to y_{3,t}$, is stationary, and can be written as:

$$\begin{aligned} y_{1,t} &= \varepsilon_{1,t} \\ y_{2,t} &= a_{21} y_{1,t-1} + \varepsilon_{2,t} \\ y_{3,t} &= a_{32} y_{2,t-1} + \varepsilon_{3,t} \ . \end{aligned}$$

Since $y_{1,t} \not\to y_{3,t}$ and $y_{3,t} \not\to y_{2,t}$, it follows from Proposition 3 that $\rho(\bar{Y}_{3,T}, \bar{Y}_{1,T}|\bar{Y}_{2,T}) = 0$, and that there is no contemporaneous causality between $\bar{Y}_{1,T}$ and $\bar{Y}_{3,T}$. The contemporaneous relationship of the aggregated limiting process may therefore be expressed graphically as $\bar{Y}_{1,T} \Rightarrow \bar{Y}_{2,T} \Rightarrow \bar{Y}_{3,T}$. According to Proposition 1, the limiting process is white noise with covariance matrix given by:

$$E(\bar{Y}_T \bar{Y}_T') = (I - A)^{-1}\Omega(I - A')^{-1}.$$

This process can be represented as:

$$(I - A)\bar{Y}_T = U_T,$$

where $E(U_T U_T') = \Omega$. Specifically,

$$\begin{aligned}
\bar{Y}_{1,T} &= U_{1,t} \\
\bar{Y}_{2,T} &= a_{21}\bar{Y}_{1,T} + U_{2,T} \\
\bar{Y}_{3,T} &= a_{32}\bar{Y}_{2,T} + U_{3,T} \ .
\end{aligned}$$

There is, however, an equivalent representation given by $\bar{Y}_{3,T} \Rightarrow \bar{Y}_{2,T} \Rightarrow \bar{Y}_{1,T}$ and

$$\begin{aligned}
\bar{Y}_{3,T} &= \tilde{U}_{1,t} \\
\bar{Y}_{2,T} &= \tilde{a}_{21}\bar{Y}_{3,T} + \tilde{U}_{2,T} \\
\bar{Y}_{1,T} &= \tilde{a}_{32}\bar{Y}_{2,T} + \tilde{U}_{3,T} \ .
\end{aligned}$$

Further, there exist other representations which obey the above partial correlation restriction. For instance, it is easy to verify that in the system:

$$\begin{aligned}
\bar{Y}_{2,T} &= U_{1,t}^{*} \\
\bar{Y}_{1,T} &= a_{21}^{*}\bar{Y}_{2,T} + U_{2,T}^{*} \\
\bar{Y}_{3,T} &= a_{32}^{*}\bar{Y}_{2,T} + U_{3,T}^{*}
\end{aligned}$$

the partial correlation restriction is fulfilled as well. This representation results from reversing the direction of the first arrow in the original graph so that $\bar{Y}_{1,T} \Leftarrow \bar{Y}_{2,T} \Rightarrow \bar{Y}_{3,T}$.

It is obvious that the possibility of different representations introduces some ambiguity to any discussion of correspondence between Granger and instantaneous causality. Thus, in order to consider the relationship between both causality concepts we therefore need to abstract from the direction of causality.[3]

**Example B:** Assume that a vector of flow variables is generated by a stationary process given by:

$$\begin{aligned} y_{1,t} &= a_{12}y_{2,t-1} + a_{13}y_{3,t-1} + \varepsilon_{1,t} \\ y_{2,t} &= \varepsilon_{2,t} \\ y_{3,t} &= \varepsilon_{3,t} \ . \end{aligned}$$

Applying Granger's concept of causality, there is no causality between $y_{2,t}$ and $y_{3,t}$. Further, a simple calculation shows that for the limiting process, $\rho(\bar{Y}_{2,T}, \bar{Y}_{3,T}|\bar{Y}_{1,T}) = -a_{12}a_{13}/(a_{12}^2 + a_{13}^2 + 1)$. Thus, a necessary and sufficient condition for the aggregated variables $\bar{Y}_{2,T}$ and $\bar{Y}_{3,T}$ to have no contemporaneous causal relationship is that either $a_{12}$, $a_{13}$, or both parameters are equal to zero. This follows from Proposition 3, which states that there is no contemporaneous causality if either $y_{2,t}$ or $y_{3,t}$ is not Granger causal for $y_{1,t}$.

**Example C:** To illustrate the problems with aggregated stock variables which are discussed above, consider the stationary process given by:

$$\begin{aligned} y_{1,t} &= \varepsilon_{1,t} \\ y_{2,t} &= a_{21}y_{1,t-1} + \varepsilon_{2,t} \\ y_{3,t} &= a_{32}y_{2,t-1} + \varepsilon_{3,t}. \end{aligned}$$

In this system, $y_{1,t} \to y_{2,t}$ and $y_{2,t} \to y_{3,t}$. For $k \geq 3$ the aggregated process becomes white noise with:

$$Y_{1,T} = U_{1,T}$$

[3]The same sort of problem arises when orthogonalizing errors in a VAR using the method proposed by SG (1997). This problem is referred to as a reversibility problem by SG, arises because $\gamma(x,y) = \gamma(y,x)$, for any two random variables, $x$ and $y$. In their context, the cost of this problem is that the group of recursive models from among which a "final" structural model of the errors can be chosen can only be narrowed down to two, at which stage economic theory may be used to "choose" a final model.

$$Y_{2,T} = U_{2,T}$$
$$Y_{3,T} = (a_{21}a_{32})Y_{1,T} + U_{3,T}$$

For $(a_{21}a_{32}) \neq 0$ there exists *spurious* contemporaneous causality between $Y_{1,T}$ and $Y_{3,T}$, as there is no Granger causality between $y_{1,t}$ and $y_{3,t}$. Stated another way, the indirect causal relationship between $y_{1,t}$ and $y_{3,t}$ via $y_{2,T}$ becomes a direct causal link (i.e. $Y_{1,T} \Rightarrow Y_{3,T}$), under aggregation.

# 6 Monte Carlo Experiments

So far we have considered the relationship between Granger causality and instantaneous causality, when the aggregation interval $k$ tends to infinity. In practical applications, however, $k$ is finite, so that the limiting process does not render the actual representation of the system, in general. It is therefore of interest to assess the ability of our asymptotic results to provide useful guidance for finite aggregation intervals. For the sake of simplicity, we focus on a VAR(1) model of the form examined by Swanson (1994). The data generating process is given by:

$$\begin{bmatrix} \Delta^d y_{1,t} \\ \Delta^d y_{2,t} \\ \Delta^d y_{3,t} \end{bmatrix} = \begin{bmatrix} a & 0 & 0 \\ b & a & 0 \\ 0 & b & a \end{bmatrix} \begin{bmatrix} \Delta^d y_{1,t-1} \\ \Delta^d y_{2,t-1} \\ \Delta^d y_{3,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{bmatrix}, \tag{14}$$

where the $\varepsilon_{i,t}$ is an i.i.d. vector of standard normal random variables. For $b \neq 0$, the Granger causal structure of this system is: $y_{1,t} \to y_{2,t} \to y_{3,t}$. From Proposition 3 and Corollary 1 it follows that, as $k$ tends to infinity, the limiting process has a corresponding contemporaneous causal structure, whenever $y_t$ is a vector of (i) stationary flow variables, (ii) difference stationary flow variables, or (iii) difference stationary stock variables. In all of these cases the empirical procedure suggested by Swanson and Granger (1997) should indicate the correct contemporaneous causal relationship, given Proposition 3 above.

Let $\widehat{U}_{i,T}$ denote the residuals from a VAR($p$) regression of $Y_T$ or $\bar{Y}_T$. SG (1997) propose specifying the causal structure by testing all whether partial correlations of that form $\rho(\widehat{U}_{i,T}, \widehat{U}_{j,T}|\widehat{U}_{l,T})$ which are implied by an assumed causal ordering are zero, for $i \neq j \neq l$. Here, the empirical procedure of Swanson and Granger

**Table 1:** Probability of Selecting the Correct Causal Ordering

| a) Stationary flow variables | | | | | |
|---|---|---|---|---|---|
| k | a=0 | a=0.2 | a=0.4 | a=0.6 | a=0.8 |
| 2 | 0.899 | 0.892 | 0.887 | 0.889 | 0.893 |
| 5 | 0.932 | 0.915 | 0.840 | 0.665 | 0.391 |
| 10 | 0.932 | 0.929 | 0.899 | 0.654 | 0.045 |
| 20 | 0.933 | 0.932 | 0.923 | 0.813 | 0.015 |
| 50 | 0.931 | 0.931 | 0.932 | 0.909 | 0.203 |
| 100 | 0.929 | 0.929 | 0.929 | 0.926 | 0.534 |
| b) Difference stationary flow variables | | | | | |
| k | a=0 | a=0.2 | a=0.4 | a=0.6 | a=0.8 |
| 2 | 0.890 | 0.886 | 0.885 | 0.886 | 0.891 |
| 5 | 0.919 | 0.900 | 0.832 | 0.656 | 0.394 |
| 10 | 0.928 | 0.926 | 0.902 | 0.693 | 0.073 |
| 20 | 0.932 | 0.930 | 0.927 | 0.863 | 0.054 |
| 50 | 0.934 | 0.934 | 0.933 | 0.929 | 0.474 |
| 100 | 0.927 | 0.928 | 0.930 | 0.930 | 0.830 |

**Notes:** Entries correspond to the frequency of times that the correct contemporaneous causal structure is uncovered, based on empirical procedure given in Swanson and Granger (1997). Results are based on estimations using 100 observations of data generated according to (14), and aggregated according to the aggregation interval, $k$. All entries are based on 10,000 Monte Carlo replications.

(1997) uncovers the correct causal ordering if $\rho(\widehat{U}_{3,T}, \widehat{U}_{1,T}|\widehat{U}_{2,T}) = 0$ and all other partial correlations are different from zero[4]. In our Monte Carlo experiment we estimate the probability that a test based on $\rho(\widehat{U}_{3,T}, \widehat{U}_{1,T}|\widehat{U}_{2,T})$ fails to reject the null hypothesis: $H_0 : \rho(\widehat{U}_{3,T}, \widehat{U}_{1,T}|\widehat{U}_{2,T}) = 0$. This hypothesis is tested using a test based on the well known Fisher's $z$-statistic (e.g. see Anderson (1985)). Frequencies of failure to reject this null hypothesis can be interpreted as signalling the probability of selecting the "correct" contemporaneous causal ordering. Further, the asymptotic results of Section 3 imply that as $k \to \infty$, $\rho(\widehat{U}_{3,T}, \widehat{U}_{1,T}|\widehat{U}_{2,T})$ converges to zero, so that our experiment also provides evidence concerning the usefulness of our asymptotic results in the context of finite $k$.

We use a 5 percent significance level for our tests, so that we expect that the

[4]See SG (1997) for details concerning which partial correlations can be validly tested based on an assumed structural model of the errors

**Table 2:** Probability of Selecting the Correct Causal Ordering

| a) Stationary stock variables | | | | | |
|---|---|---|---|---|---|
| k | a=0 | a=0.2 | a=0.4 | a=0.6 | a=0.8 |
| 2 | 0.931 | 0.932 | 0.933 | 0.927 | 0.903 |
| 5 | 0.932 | 0.932 | 0.931 | 0.910 | 0.741 |
| 10 | 0.931 | 0.928 | 0.928 | 0.896 | 0.481 |
| 20 | 0.930 | 0.932 | 0.932 | 0.900 | 0.344 |
| 50 | 0.934 | 0.932 | 0.929 | 0.897 | 0.344 |
| 100 | 0.931 | 0.929 | 0.926 | 0.897 | 0.343 |
| b) Difference stationary stock variables | | | | | |
| k | a=0 | a=0.2 | a=0.4 | a=0.6 | a=0.8 |
| 2 | 0.899 | 0.892 | 0.887 | 0.889 | 0.893 |
| 5 | 0.932 | 0.915 | 0.840 | 0.665 | 0.391 |
| 10 | 0.932 | 0.929 | 0.899 | 0.654 | 0.045 |
| 20 | 0.933 | 0.932 | 0.923 | 0.813 | 0.015 |
| 50 | 0.923 | 0.931 | 0.932 | 0.909 | 0.203 |
| 100 | 0.929 | 0.929 | 0.929 | 0.926 | 0.534 |

**Notes:** See notes to Table 1.

estimated probabilities in the tables to be (approximately) 0.95. All entries in the tables are based on 10,000 Monte Carlo replications, and all estimations use 100 observations of appropriately aggregated data. (Results based on samples of 500 observations are available on request.) Also, we use VAR(4) models in order to estimate the errors of the system, although our results are not sensitive to lag order. Not surprisingly, the magnitude of the parameter $a$ is crucial to the applicability of our asymptotic results, when $k$ is small. In particular, as $a$ determines the roots of the autoregressive polynomial in our model, our asymptotic results may be a poor guide to finite sample behavior, when $k$ is small and the absolute value of $a$ is close to unity. Given this consideration, we allow $a$ to take a range of values (i.e. $a$=$\{0.0, 0.2, 0.4, 0.6, 0.8\}$). In contrast, the results are not very sensitive to the parameter $b$. For the sake of brevity, we therfore fix $b = 0.5$.

Tables 1 and 2 present the frequencies of selecting the correct contemporaneous causal ordering, for various parameter values. [5] Table 1 a) gives results

[5]It is worth reiterating that the aggregated processes which we construct are VARMA processes, in general. Thus, lower order VAR approximations may not yield good estimates of the

for aggregated flow variables. It turns out that the partial correlation dies out quickly as $a$ approaches zero and, thus, the rejection frequencies are close to 0.95. For large values of $a$, however, there is a sizeable partial correlation for moderate values of $k$, and the probability of selecting the correct contemporaneous causal ordering converges quite slowly to the asymptotic value of 0.95. For difference stationary flow variables, our findings are qualitatively similar (see Table 1 b).

Based on our asymptotic results for *stationary stock variables*, we know that contemporaneous causality need not correspond to Granger causality as $k \to \infty$. For $a = 0$, however, we have $A^n = 0$ for $n \geq 3$ and the covariance matrix of the limiting process is $E(Y_T Y_T') = \Omega + A\Omega A' + A^2\Omega(A^2)'$ (see Proposition 1 (i)), which is a diagonal matrix in this special case. Therefore, all partial correlations are zero, and *spurious* contemporaneous correlation should not arise for $a = 0$. In contrast, for $a$ substantially different from zero, partial correlations need not die out as $k$ increases. In fact for $a = 0.8$ the frequency of selecting the correct causal ordering is quite different from 0.95 for moderate and large values of $k$. For small values of $k$, however, the partial correlations is small, although the SG (1997) show some evidence of bias in these cases.

For the case when aggregated *stock variables* are generated by a *difference stationary* process (see Table 2 b), partial correlations should converge to zero as $k \to \infty$, according to Proposition 3. The simulation results clearly reveal this property, even for relatively small values of $k$, although the frequency of times that the correct contemporaneous causal ordering is selected converges rather slowly to 0.95, as $k$ increases. Summarizing our experimental findings, it appears that our asymptotic results are applicable for small values of $k$, within the context of the simple VAR model which we examine.

errors of the process. Second, even if correct VARMA models are estimated, the error covariance matrices are generally different from the asymptotic covariance matrices corresponding to $k = \infty$. Finally, the test we use has an *asymptotic* significance level of 0.05, whereas in small samples, the actual size may be different. In fact, our simulations indicate that there is a moderate size bias if the partial correlation is zero (e.g. for $a = 0$).

# 7 Concluding Remarks

In this paper we examine the effects of temporal aggregation on stock and flow data within the context of characterizing Granger causal and contemporaneous relationships among systems of economic variables which are examined using VAR models. Usually, the sampling frequency of the data is different from the original frequency and it is thus important to investigate whether it is possible to infer (Granger) causality of the original process from the properties of the aggregated process. However, to obtain analytical results as in Christiano and Eichengreen (1987) the aggregation interval must be given, which is usually not known in practice. In this paper we derive asymptotic results by assuming that the aggregation interval is large relative to the original time scale of the underlying data generating process.

Under some plausible conditions it can be shown that there is some kind of correspondence between the concepts of Granger and contemporaneous causality. Hence, we are able to make statements about the underlying causal structure by considering highly aggregated data. Nevertheless, there is an obvious loss of information when considering contemporaneous causality. Since aggregation affect the temporal ordering of cause and effect, the direction of causality is not identified. Hence, economic theory must be employed to infer the direction of causality.

Our analysis may also be useful to motivate structural versions of vector autoregression as suggested by Sims (1986) and Blanchard and Watson (1986). The (Granger) causal structure of the original process implies a set of corresponding relationships among the innovation of the process. Accordingly, the structural identification of the shocks should mirror the underlying causal structure of the multivariate process.

# Appendix: Proofs

## Proposition 1:

(i) From $Y_T = y_{kT}$ and assuming stationarity we have

$$E(Y_T Y_T') = E(y_t y_t') = \left(I_m + \sum_{i=1}^{\infty} C_i\right) \Omega \left(I_m + \sum_{i=1}^{\infty} C_i'\right)$$

(ii) Since the process is assumed to be ergodic we have

$$\lim_{k\to\infty} E(Y_T Y_{T+j}') = \lim_{k\to\infty} E(y_{kT} y_{kT+jk}') = 0$$

for all $j \geq 1$.

(iii) The vector of aggregated flow variables is given by

$$\bar{Y}_T = \sum_{j=0}^{k-1} y_{kT-j}$$

and therefore $\bar{Y}_T$ behaves as a vector partial sum. For partial sums it is known that

$$\lim_{k\to\infty} E(\frac{1}{k}\bar{Y}_T \bar{Y}_T') = \Omega + \Gamma + \Gamma',$$

where $\Gamma = \sum_{j=1}^{\infty} E(y_t y_{t-j}')$. In the frequency domain this expression can be represented as

$$2\pi f_y(0) = \left(I_m - \sum_{j=1}^{\infty} A_j\right)^{-1} \Omega \left(I_m - \sum_{j=1}^{\infty} A_j'\right)^{-1}.$$

(iv) Let

$$\begin{aligned}
\bar{Y}_T &= (I_m + L + L^2 + L^{k-1})C(L)\varepsilon_t \\
&\equiv D(L)\varepsilon_t,
\end{aligned}$$

where

$$D(L) = I_m + D_1 L + D_2 L^2 + \cdots$$

and

$$D_j = \sum_{i=0}^{\min(j,k-1)} C_{j-i}.$$

It is convenient to decompose $\bar{Y}_T$ as

$$\begin{aligned}
\bar{Y}_T &= D_0\varepsilon_t + D_k\varepsilon_{t-k} + D_{2k}\varepsilon_{t-2k} + \cdots \\
&\quad +D_1\varepsilon_{t-1} + D_{k+1}\varepsilon_{t-k-1} + D_{2k+1}\varepsilon_{t-2k-1} + \cdots \\
&\quad \vdots \\
&\quad +D_{k-1}\varepsilon_{t-k+1} + D_{2k-1}\varepsilon_{t-2k+1} + D_{3k-1}\varepsilon_{t-3k+1} + \cdots \\
&\equiv u_{0t} + \cdots + u_{k-1,t}
\end{aligned}$$

where

$$u_{jt} = D_j\varepsilon_{t-j} + D_{j+k}\varepsilon_{t-j-k} + \cdots$$

Note that $E(u_{it}u'_{jt}) = 0$ for $i \neq j$.

From

$$\begin{aligned}
\bar{Y}_T &= u_{0t} + \cdots + u_{k-1,t} \\
\bar{Y}_{T+1} &= u_{0,t+k} + \cdots + u_{k-1,t+k} \\
\bar{Y}_{T+2} &= u_{0,t+2k} + \cdots + u_{k-1,t+2k}
\end{aligned}$$

we obtain:

$$E(\bar{Y}_T\bar{Y}'_{T+1}) = \sum_{j=0}^{k-1} E(u_{jt}u'_{j,t+k}).$$

Consider

$$E(u_{0t}u'_{0,t+k}) = D_0\Omega D'_k + D_k\Omega D'_{2k} + \cdots.$$

For a summable sequence $C_i$ we have

$$\lim_{k\to\infty} |D_{2k}| = \lim_{k\to\infty} |C_{k+1} + C_{k+2} + \cdots + C_{2k}| = 0$$

so that

$$\begin{aligned}
\lim_{k\to\infty} E(u_{0t}u'_{0,t+k}) &= D_0\Omega D'_k \\
&= \Omega(C_1 + C_2 + \cdots + C_k)'.
\end{aligned}$$

Similarly we get:

$$\begin{aligned}
\lim_{k\to\infty} E(u_{1t}u'_{1,t+k}) &= D_1\Omega D'_{k+1} \\
&= (I_m + C_1)\Omega(C_2 + C_3 + \cdots + C_{k+1})' \\
\lim_{k\to\infty} E(u_{k-1,t}u'_{k-1,t-k}) &= (C_1 + \cdots + C_{k-1})\Omega(C_k + C_{k+1} + \cdots + C_{2k-1})'.
\end{aligned}$$

Adding these expressions gives the desired result.

It remains to show that $\sum_j (\sum_{i=0}^{j} C_i)\Omega(\sum_{i=j+1}^{\infty} C_i)'$ is bounded.

Let $\bar{c} = \sup_t \left\| \sum_{j=0}^{t} C_j \right\|$.

Then:

$$\left\| \sum_j \left( \sum_{i=0}^{j} C_i \right) \Omega \left( \sum_{i=j+1}^{\infty} C'_i \right) \right\| \leq \sum_j \left\| \sum_{i=0}^{j} C_i \right\| \|\Omega\| \sum_{i=j+1}^{\infty} j \|C_i\|$$

which is finite by assumption.

(iii) Consider

$$E(u_{0t}u'_{0,t-pk}) = D_0\Omega D'_{pk} + D_k\Omega D_{(p+1)k} + \cdots.$$

Since

$$\lim_{k\to\infty} D_{(p+j)k} = 0 \qquad \text{for } p \geq 2 \text{ and } j = 0, 1, \ldots$$

it follows that the autocovariances disappear for $p \geq 2$.

## Proposition 2

**(i)** The difference

$$Y_T - Y_{T-1} = y_{kT} - y_{kT-k} = \sum_{i=1}^{k} \Delta y_{(k-1)T+i}$$

is a partial sum process with asymptotic covariance matrix

$$\begin{aligned}
\lim_{k\to\infty} k^{-1} E(Y_T - Y_{T-1})(Y_T - Y_{T-1})' &= \Omega + \Gamma + \Gamma' \\
&= 2\pi f_{\Delta y}(0)
\end{aligned}$$

**(ii)** Define the partial sum $S_1 = \sum_{i=1}^{k} u_i$ and $S_2 = \sum_{i=k+1}^{2k} u_i$, where $u_t$ is stationary with covariance function $\Gamma_j$. The covariance between $S_1$ and $S_2$ is given by to

$$E(S_1 S_2') = \Gamma_1 + 2\Gamma_2 + \cdots + k\Gamma_k + (k-1)\Gamma_{k+1} + \cdots + \Gamma_{2k-1}$$

For $\sum_{j=1}^{\infty} j|\Gamma_j| < \infty$ we have

$$\begin{aligned} |E(S_1 S_2')| &< |\sum_{j=1}^{\infty} j\Gamma_j| \\ &\leq \sum_{j=1}^{\infty} j|\Gamma_j| < \infty \end{aligned}$$

and, thus, by letting $S_1 = Y_T - Y_{T-1}$ and $S_2 = Y_{T+1} - Y_T$ it follows that $E(Y_T - Y_{T-1})(Y_{T+1} - Y_T)$ is $O(1)$. A similar result is obtained for higher order autocovariances.

**(iii)** Let

$$\begin{aligned} \bar{Y}_T - \bar{Y}_{T-1} &= y_{kT} - y_{kT-k} + y_{kT-1} - y_{kT-k-1} + \cdots + y_{kT-k+1} - y_{kT-2k+1} \\ &= S_k(L)\Delta y_{kT} + S_k(L)\Delta y_{kT-1} + \cdots + S_k(L)\Delta y_{kT-k+1} \\ &= S_k(L)^2 \Delta y_{kT}, \end{aligned}$$

where

$$S_k(L) = 1 + L + L^2 + \cdots + L^{k-1}$$

and

$$\begin{aligned} S_k(L)^2 &= 1 + 2L + 3L^2 + \cdots + kL^{k-1} + (k-1)L^k + \cdots + L^{2k-2} \\ &= w_0 + w_1 L + w_2 L + \cdots + w_{2k-2} L^{2k-2} \end{aligned}$$

is a symmetric filter with triangular weights.

The covariance matrix is given by

$$\begin{aligned} E(\bar{Y}_T - \bar{Y}_{T-1})(\bar{Y}_T - \bar{Y}_{T-1})' &= E\left(\sum_{i=0}^{2k-2} w_i \Delta y_{kT-i}\right)\left(\sum_{i=0}^{2k-2} w_i \Delta y'_{kT-i}\right) \\ &= \sum_{p=-2k+2}^{2k-2} \sum_{i=1}^{2k-1-|p|} w_i w_{i+|p|} \Gamma_p \end{aligned}$$

where $\Gamma_p = E(\Delta y_t \Delta y'_{t-p})$.

Consider the odd values $p = \pm 1, \pm 3, \pm 5, \ldots$. We have

$$\sum_{i=1}^{2k-1-|p|} w_i w_{i+|p|} = 2 \sum_{i=1}^{k-(|p|+1)/2} i(i+p)$$

and as $k \to \infty$

$$\begin{aligned}\lim_{k\to\infty} \sum_{i=1}^{k-|(p+1)/2|} 2(i^2 - ip) &= 2(\sum_{i=1}^{\infty} i^2) - 2p(\sum_{i=1}^{\infty} i) \\ &= \frac{2}{3}k^3 + O(k^2).\end{aligned}$$

For even values $p = 0, \pm 2, \pm 4, \ldots$ we have

$$\sum_{i=1}^{2k-1-|p|} w_i w_{i+|p|} = (k - |p|/2)^2 + 2 \sum_{i=1}^{k-|p|/2-1} i(i+p)$$

and, thus,

$$\lim_{k\to\infty} \sum_{i=1}^{k-|(p+1)/2|} 2(i^2 - ip) = \frac{2}{3}k^3 + O(k^2).$$

Using these results yields

$$\begin{aligned}E(\bar{Y}_T - \bar{Y}_{T-1})(\bar{Y}_T - \bar{Y}_{T-1})' &= \frac{2}{3}k^3(\Gamma_0 + \sum_{j=1}^{\infty} \Gamma_j + \Gamma'_j) + o(k^3) \\ &= \frac{4\pi}{3}k^3 f_{\Delta y}(0) + o(k^3).\end{aligned}$$

**(iv)** The first order autocovariance matric is given by

$$E(\bar{Y}_T - \bar{Y}_{T-1})(\bar{Y}_{T+1} - \bar{Y}_T)' = \sum_{p=-2k+2}^{2k-2} \sum_{i=1}^{2k-1-|p|} w_{i+k} w_{i+k+|p|} \Gamma_p$$

where $\Gamma_p = E(\Delta y_t \Delta y'_{t-p})$.

For an odd value of $p$ we have

$$\begin{aligned}\lim_{k\to\infty} \sum_{i=1}^{2k-1-|p|} w_{i+k} w_{i+k+|p|} &= \sum_{i=1}^{\infty} (k-i)(i+p) + O(k^2) \\ &= k(\sum_{i=1}^{\infty} i) - (\sum_{i=1}^{\infty} i^2) + k^2 p - p(\sum_{i=1}^{\infty} i) + O(k^2) \\ &= \frac{1}{6}k^3 + O(k^2)\end{aligned}$$

It follows that

$$\begin{aligned} E(\bar{Y}_T - \bar{Y}_{T-1})(\bar{Y}_{T+1} - \bar{Y}_T)' &= \frac{1}{6}k^3(\Gamma_0 + \sum_{j=1}^{\infty}\Gamma_j + \Gamma_j') \\ &= \frac{\pi}{3}k^3 f_{\Delta y}(0). \end{aligned}$$

**(v)** To simplify the proof we assume that $\Delta y_t$ has an vector MA($q$) representation with $q < k$. Since $k \to \infty$ the proof is valid for $q \to \infty$, as well. Of course, the assumption $q < k$ imposes the restriction that $q$ does not grow at a faster rate that $k$. This assumption is not necessary for the proof. However, a more general treatment would complicate the proof substantially.

The second order autocovariance matrix is given by

$$\begin{aligned} E(\bar{Y}_T - \bar{Y}_{T-1})(\bar{Y}_{T+2} - \bar{Y}_{T+1})' &= E\left(\sum_{i=0}^{2k-2} w_i \Delta y_{kT-i}\right)\left(\sum_{i=0}^{2k-2} w_i \Delta y'_{kT+2k-i}\right) \\ &= \sum_{p=1}^{k}\sum_{i=1}^{|p|} w_i w_{2k-i-|p|+1}(\Gamma_p + \Gamma_p') \end{aligned}$$

There exist a constant $c < \infty$ such that for all $p$

$$\sum_{i=1}^{|p|} w_i w_{2k-i-|p|+1} = \sum_{i=1}^{p} i(p-i+1) < cp^3.$$

Thus, we it follows

$$\begin{aligned} \sum_{p=1}^{k}\sum_{i=1}^{|p|} w_i w_{2k-i-|p|+1}|\Gamma_p + \Gamma_p'| &< \sum_{p=1}^{k} 2cp^3|\Gamma_p| \\ &< 2ck^2\sum_{p=1}^{k} p|\Gamma_p|, \end{aligned}$$

where we have used $p < k$. From $\sum_{p=1}^{k} p|\Gamma_p| < \infty$ it finally follows that

$$\lim_{k\to\infty}\frac{1}{k^3}E(\bar{Y}_T - \bar{Y}_{T-1})(\bar{Y}_{T+2} - Y_{T+1})' = 0.$$

Similarly it can be shown that the higher order autocorrelations converge to zero as well.

## Proposition 3

For convenience, we confine ourselves to a trivatiate VAR($p$) process. The proof can easily be generalized to systems with $m > 3$.

First consider a VAR process obeying the conditions:

$$\begin{aligned} & y_{1,t} \not\to y_{2,t} \\ & y_{2,t} \not\to y_{1,t} \\ (a) \quad & y_{1,t} \not\to y_{3,t}, \end{aligned}$$

that is, there is no causality between $y_{1,t}$ and $y_{2,t}$ and condition (a) is satisfied. As $k \to \infty$ the aggregated process for stationary flow variables takes the form

$$\begin{bmatrix} \bar{Y}_{1,T} \\ \bar{Y}_{2,T} \\ \bar{Y}_{3,T} \end{bmatrix} = \begin{bmatrix} \bar{a}_{11} & 0 & \bar{a}_{13} \\ 0 & \bar{a}_{22} & \bar{a}_{23} \\ 0 & \bar{a}_{32} & \bar{a}_{33} \end{bmatrix} \begin{bmatrix} \bar{Y}_{1,T} \\ \bar{Y}_{2,T} \\ \bar{Y}_{3,T} \end{bmatrix} + \begin{bmatrix} U_{1,T} \\ U_{2,T} \\ U_{3,T} \end{bmatrix}$$

or

$$\bar{Y}_T = \bar{A} Y_T + U_T,$$

where $\bar{A} = \sum_{i=1}^{p} A_i$, $E(U_T U_T') = \Omega$ and

$$E(Y_T Y_T') = (I - \bar{A})^{-1} \Omega (I - \bar{A}')^{-1}.$$

There exists a process

$$\begin{bmatrix} \bar{Y}_{1,T} \\ \bar{Y}_{2,T} \\ \bar{Y}_{3,T} \end{bmatrix} = \begin{bmatrix} \bar{a}_{11} & 0 & \bar{a}_{13} \\ 0 & \bar{a}_{22}^{*} & \bar{a}_{23}^{*} \\ 0 & 0 & \bar{a}_{33}^{*} \end{bmatrix} \begin{bmatrix} \bar{Y}_{1,T} \\ \bar{Y}_{2,T} \\ \bar{Y}_{3,T} \end{bmatrix} + \begin{bmatrix} U_{1,T} \\ U_{2,T}^{*} \\ U_{3,T}^{*} \end{bmatrix}$$

or

$$\bar{Y}_T = \bar{A}^{*} Y_T + U_T^{*},$$

such that $E(U_T^{*} U_T^{*\prime}) = \Omega$ and

$$(I - \bar{A})^{-1} \Omega (I - \bar{A}')^{-1} = (I - \bar{A}^{*})^{-1} \Omega (I - \bar{A}^{*\prime})^{-1},$$

so that the limiting process of $Y_T$ and $Y_T^{*}$ are identical. The process $Y_T^{*}$ can be found from a Choleski decomposition of the lower-right $2 \times 2$ block of $\bar{A}$. Since the first column of $\bar{A}$ and $\bar{A}^{*}$ is identical, it follows that $\gamma(Y_{1,T}, Y_{2,T} | Y_{3,T})$ is proportional to $a_{12}^{*} = a_{12} = 0$.

Second, consider the the condition $y_{3,t} \not\to y_{1,t}$ instead of $y_{1,t} \not\to y_{3,t}$. In this case we arrange the vector $\bar{Y}_T$ as

$$\bar{Y}_T = \begin{bmatrix} \bar{Y}_{1,T} \\ \bar{Y}_{3,T} \\ \bar{Y}_{2,T} \end{bmatrix} = \begin{bmatrix} \bar{a}_{11} & \bar{a}_{12} & 0 \\ 0 & \bar{a}_{22} & \bar{a}_{23} \\ 0 & \bar{a}_{32} & \bar{a}_{33} \end{bmatrix} \begin{bmatrix} \bar{Y}_{1,T} \\ \bar{Y}_{3,T} \\ \bar{Y}_{2,T} \end{bmatrix} + \begin{bmatrix} U_{1,T} \\ U_{2,T} \\ U_{3,T} \end{bmatrix}$$

Again we may use a Choleski decomposition to find an equivalent process such that

$$\begin{bmatrix} \bar{Y}_{1,T} \\ \bar{Y}_{3,T} \\ \bar{Y}_{2,T} \end{bmatrix} = \begin{bmatrix} \bar{a}_{11} & \bar{a}_{12} & 0 \\ 0 & \bar{a}^*_{22} & \bar{a}^*_{23} \\ 0 & 0 & \bar{a}^*_{33} \end{bmatrix} \begin{bmatrix} \bar{Y}_{1,T} \\ \bar{Y}_{3,T} \\ \bar{Y}_{2,T} \end{bmatrix} + \begin{bmatrix} U_{1,T} \\ U^*_{2,T} \\ U^*_{3,T} \end{bmatrix}$$

It follows that $\gamma(Y_{1,T}, Y_{2,T}|Y_{3,T})$ is proportional to $a^*_{13} = a_{13} = 0$ and, thus, $\rho(Y_{1,T}, Y_{2,T}|Y_{3,T}) = 0$.

To generalize the proof to the case $k > 3$, we let $\bar{Y}_{3,T}$ be an $(m-2)$-vector. The reasoning of the proof applies to this case in a straightforward manner.

## Corollary 2

Let $z_t = \Delta^d y_t$ and consider the AR representation of the system

$$z_t = A_1 z_{t-1} + A_2 z_{t-2} + \cdots + \varepsilon_t$$

Causality at horizon $h$ can be deduced from the conditional expectation

$$E(z_t|z_{t-h}, z_{t-h-1}, \ldots) = \pi_1^{(h)} y_{t-h} + \pi_2^{(h)} y_{t-h-1} + \pi_3^{(h)} y_{t-h-2} + \cdots$$

where the matrices $\pi_j^{(h)}$ are given in Dufour and Renault (1998). Non-causality at any horizon $y_{i,t} \not\to_\infty y_{j,t}$ implies

$$e_i' \pi_j^{(h)} e_j = 0 \quad \text{for } j = 1, 2, \ldots$$

Using the MA representation

$$z_t = \varepsilon_t + B_1 \varepsilon_{t-1} + B_2 \varepsilon_{t-2} + \cdots$$

it is not difficult to see that $\pi_1^{(h)} = B_h$ and, thus, noncausality at all horizons implies

$$e_i' B_h e_j = 0 \quad \text{for } h = 1, 2, \ldots \tag{15}$$

Assuming stationary flow variables it follows from Proposition 1 that the limiting process can be represented as $(I_3 - \bar{A})\bar{Y}_T = U_T$. From $y_{2,t} \not\rightarrow y_{3,t}$ it follows that the $(3,1)$ elements of the matrices $A_k$, $k = 1, 2, \ldots$ are zero. Accordingly, the limiting distribution can be represented as

$$(1 - \bar{a}_{11})\bar{Y}_{1,T} = \bar{a}_{12}\bar{Y}_{2,T} + \bar{a}_{13}\bar{Y}_{3,T} + U_{1,T} \tag{16}$$

$$(1 - \bar{a}_{22})\bar{Y}_{2,T} = \bar{a}_{21}\bar{Y}_{1,T} + \bar{a}_{23}\bar{Y}_{3,T} + U_{2,T} \tag{17}$$

$$(1 - \bar{a}_{33})\bar{Y}_{3,T} = \bar{a}_{31}\bar{Y}_{1,T} + U_{3,T} \tag{18}$$

where $\bar{a}_{ij}$ denotes the $(i,j)$ element of the matrix $\bar{A} = \sum A_k$. From the MA representation we get $(I - \bar{A})^{-1} = (I - \bar{B})$, where $\bar{B} = \sum B_k$. From (15) it follows that for $y_{2,t} \not\rightarrow_\infty y_{3,t}$ we have $\bar{b}_{32} = 0$. Thus, the representation $\bar{Y}_T = (I_3 - \bar{A})^{-1}U_T = (I_3 - \bar{B})U_T$ gives rise to the equation

$$\bar{Y}_{3,T} = -\bar{b}_{31}U_{1,T} + (1 - \bar{b}_{33})U_{3,T} \ .$$

Solving this equation for $U_{3,T}$ and inserting in (18) gives gives

$$\bar{Y}_{1,T} = c_1\bar{Y}_{3,T} + c_2U_{1,T} \ , \tag{19}$$

where $c_1$ and $c_2$ are functions of $\bar{a}_{31}, \bar{a}_{33}, \bar{b}_{31}$ and $\bar{b}_{33}$. Comparing (19 with (16) shows that $\bar{a}_{12}$ must be zero and, thus, implies the same restriction for the limiting process as the assumption that $y_{2,t}$ does not cause $y_{1,t}$. From Proposition 3 it follows that in this case there is no spurious contemporaneous causality between $\bar{Y}_{2,T}$ and $\bar{Y}_{3,T}$.

The proofs for the cases (i) and (iii) are essentially the same.

# 8 References

**Amemiya, T. and R.Y. Wu** (1972), The Effect of Aggregation on Prediction in the Autoregressive Model, *Journal of the American Statistical Association*, No. 339, 628-632.

**Anderson, T.W.** (1985), *An Introduction to Multivariate Statistical Analysis*, (2nd ed.), New York, Wiley.

**Brewer, K.R.W.** (1973), Some Consequences of Temporal Aggregation and Systematic Sampling for ARMA and ARMAX Models, *Journal of Econometrics*, 1, 133-154.

**Christiano and Eichenbaum** (1987), Temporal Aggregation and Structural Inference in Macroeconomics, *Carnegie-Rochester Conference Series on Public Policy*, 26, 63–130.

**Dufour, J.-M. and E. Renault** (1998), Short Run and Long Run Causality in Time Series: Theory, forthcoming in: *Econometrica*

**Engle, R.F.** (1969), Biases from Time-Aggregation of Distributed Lag Models, Ph.D. Dissertation, Cornell University.

**Engle, R.F. and C.W.J. Granger** (1987), Co-integration and Error Correction: Representation, Estimation, and Testing, *Econometrica*, 55, 251-276.

**Geweke, J.** (1986), Measurement of Linear Dependence and Feedback Between Multiple Time Series, *Journal of American Statistical Association,* 77, 304–313.

**Granger, C.W.J.** (1969), Investigating Causal Relations by Econometric Models and Cross Spectral Methods, *Econometrica*, 37, 428-438.

**Granger, C.W.J.** (1980), Testing for Causality, A Personal Viewpoint, *Journal of Economic Dynamics and Control*, 2, 329-352.

**Granger, C.W.J.** (1987), Implications of Aggregation with Common Factors, *Econometric Theory*, 32, 208-222.

**Granger, C.W.J.** (1988), Some Recent Developments in a Concept of Causality, *Journal of Econometrics*, 39, 199-211.

**Granger, C.W.J., and J.-L. Lin** (1995), Causality in the Long Run, *Econometric Theory*, 11, 530-536.

**Glymour, C., R. Scheines, P. Spirtes, and K. Kelly** (1987), *Discovering Causal Structure*, San Diego, Academic Press.

**Lütkepohl, H.** (1982), Non-Causality due to Omitted Variables, *Journal of Econometrics*, 19, 367–378.

**Lütkepohl, H.** (1987), *Forecasting Aggregated Vector ARMA Processes*, Berlin, Springer.

**Lütkepohl, H.** (1991), *Introduction to Multiple Time Series Analysis*, Berlin, Springer.

**Lütkepohl, H. and M.M. Burda** (1997), Modified Wald Tests Under Nonregular Conditions, *Journal of Econometrics*, 78, 315–332.

**Massimiliano, M.** (1996), Some Temporal Aggregation Issues in Empirical Analysis, mimeo, European University Institute.

**Pierce, D.A. and L.D. Haugh** (1977), Causality in Temporal Systems: Characterizations and a survey, *Journal of Econometrics*, 5, 265–293.

**Stram, D.O. and W.W.S. Wei** (1986), Temporal Aggregation in the ARIMA Process, *Journal of Time Series Analysis*, 7, 279-292.

**Swanson, N.R.** (1994), Essays in Forecasting Stationary and Nonstationary Stochastic Processes, Ph.D. Dissertation, University of California, San Diego.

**Swanson, N.R. and C.W.J. Granger** (1997), Impulse Response Functions Based on a Causal Approach to Residual Orthogonaliztion in Vector Autoregressions, *Journal of the American Statistical Association*, 92, 357-367.

**Wei, W.W.S.** (1978), The Effect of Temporal Aggregation on Parameter Estimation in Distributed Lag Models, *Journal of Econometrics*, 8, 237-246.

**Wei, W.W.S.** (1981), Effect of Systematic Sampling on ARIMA Models, *Communications in Statistical-Theoretical Mathematics A*, 10, 2389-2398.

**Weiss, A.A.** (1984), Systematic Sampling and Temporal Aggregation in Time Series Models, *Journal of Econometrics*, 26, 271-281.

**Working, H.** (1960), Note on the Correlation of First Differences of Averages in a Random Chain, *Econometrica*, 26, 916–918.

**Zellner, A. and C. Montmarquette** (1971), A Study of Some Aspects of Temporal Aggregation Problems in Econometric Analyses, *Review of Economics and Statistics*, 335-342.